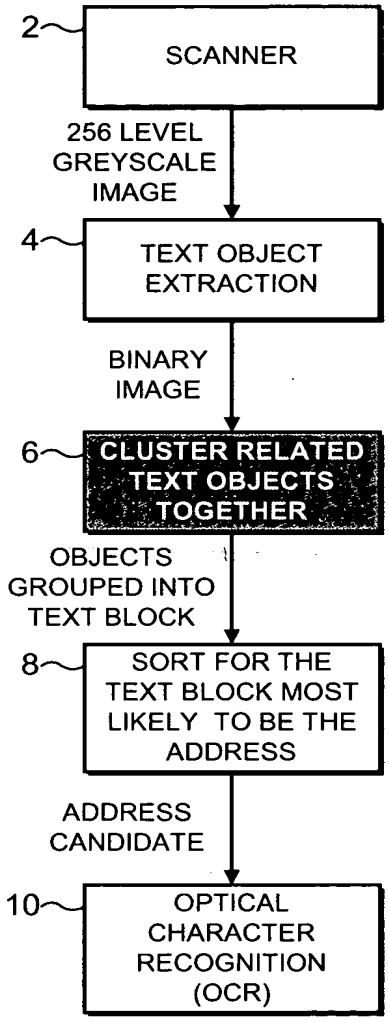

FIG. 1

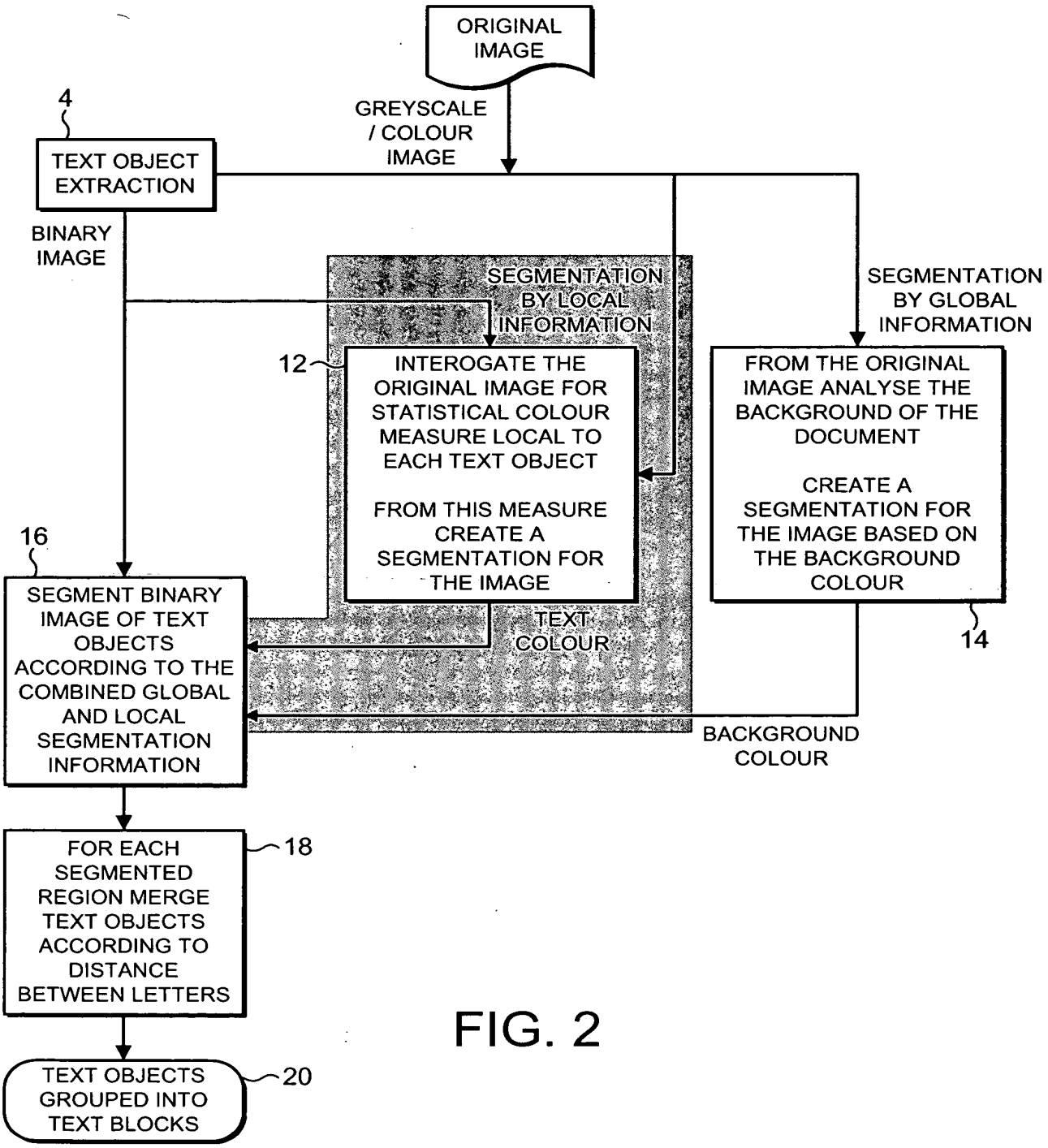

FIG. 2

A) ORIGINAL DOCUMENT IMAGE

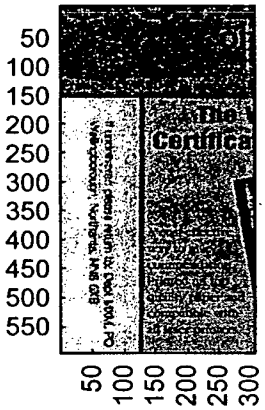

FIG. 3A

B) EXTRACTED BINARY TEXT OBJECTS

FIG. 3B

C) CLUSTERING USING SIMPLE MERGING. ALL TEXT OBJECTS ARE CLUSTERED TOGETHER

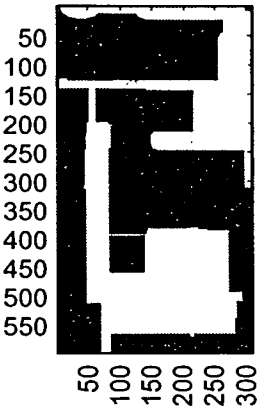

FIG. 3C

D) EXTRACT BACKGROUND FROM ORIGINAL

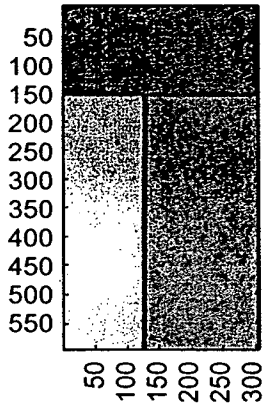

FIG. 3D

E) SEGMENT ACCORDING TO BACKGROUND. FOR EACH BACKGROUND REGION SEPARATELY APPLY MERGING TO THE TEXT OBJECTS

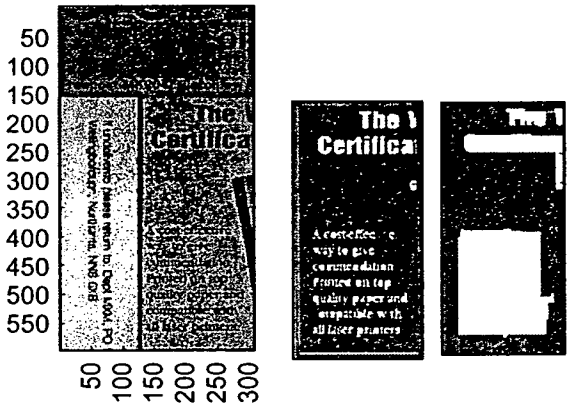

FIG. 3E

F) THE RESULTS ARE CLUSTERED TEXT OBJECTS THAT HAVE CONSISTENT BACKGROUND

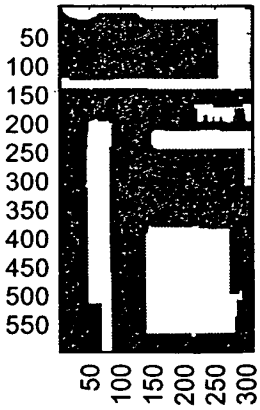

FIG. 3F

ORIGINAL DOCUMENT IMAGE

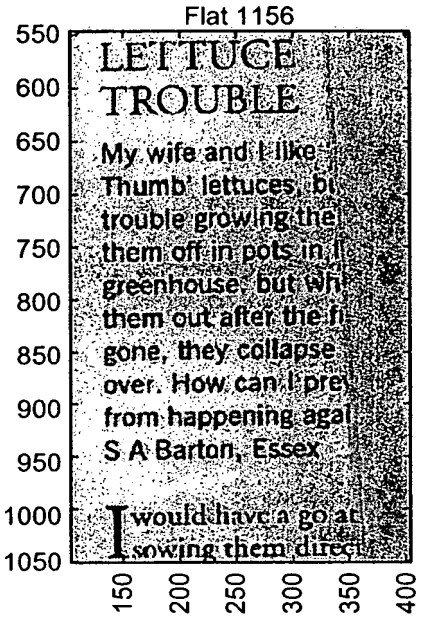

FIG. 4A

FOR EACH OF THE BINARY TEXT OBJECTS EXTRACTED THE LOCAL MINIMUM GREYLEVEL IS OBTAINED FROM THE ORIGINAL IMAGE. THE LOCAL MINIMUM GREYLEVEL IS A MEASURE OF THE TEXT COLOUR

MINIMUM GREY LEVEL IMAGE

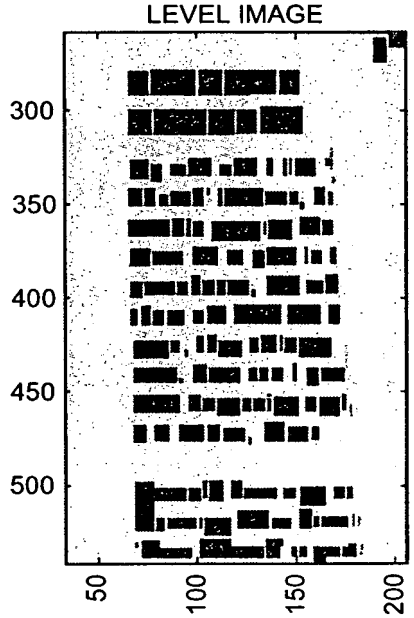

FIG. 4B

THE MINIMUM GREYLEVEL DATA IS USED TO BUILD A SEGMENTATION OF THE IMAGE. EACH REGION IN SEGMENTATION IS AN AREA OF THE IMAGE WHERE THE TEXT COLOUR IS CONSISTENT.

MERGED TEXT BLOCKS

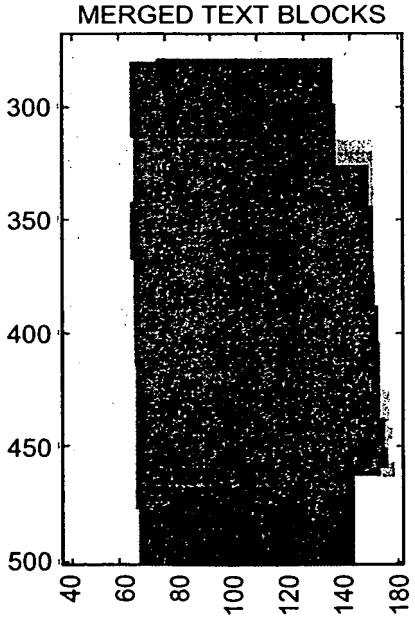

FIG. 4C

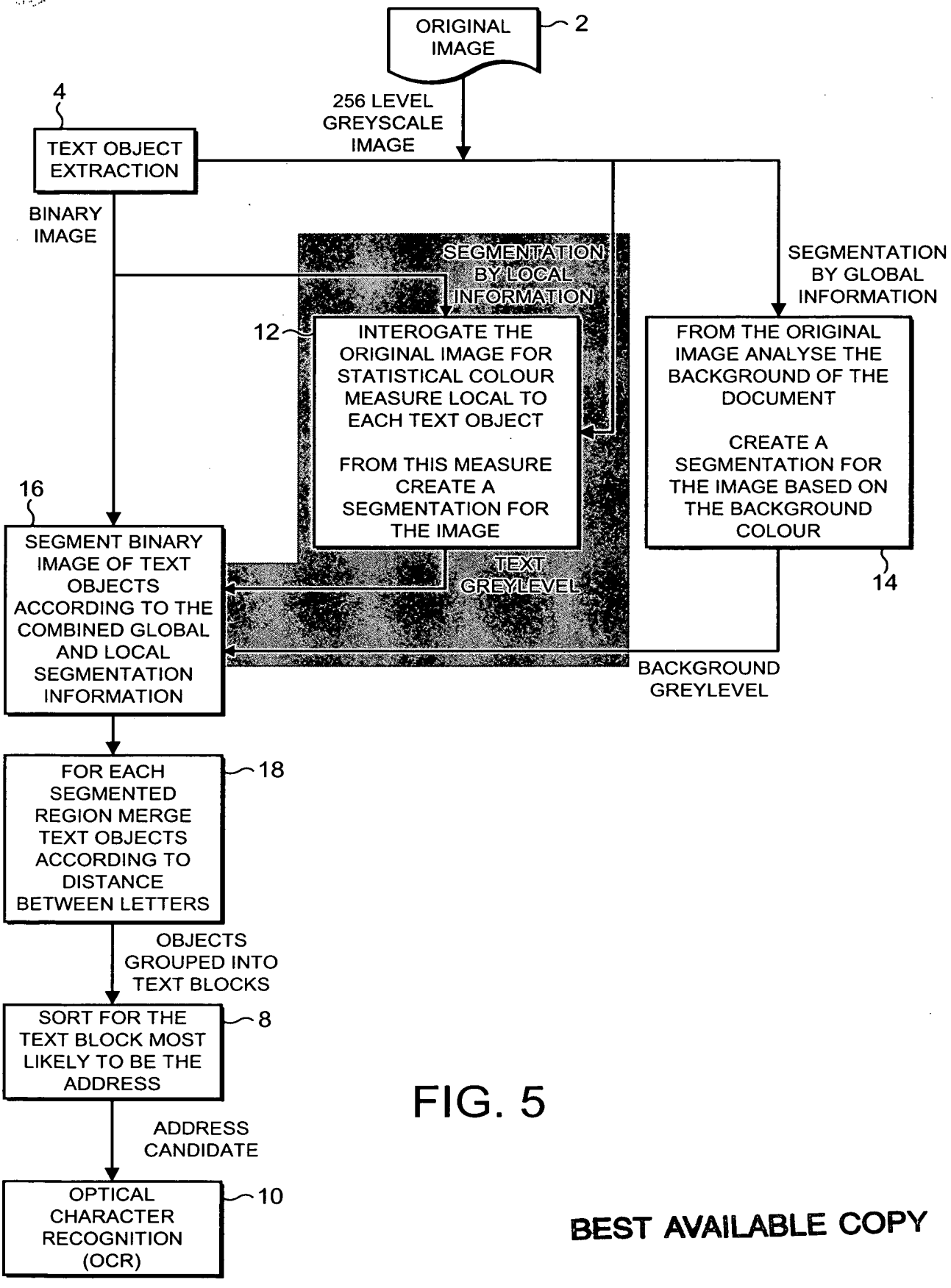

FIG. 5